(12) UK Patent Application (19) GB (11) 2 169 605 A

(43) Application published 16 Jul 1986

(21) Application No 8530915

(22) Date of filing 16 Dec 1985

(30) Priority data

(31) 59/281645 (32) 28 Dec 1984 (33) JP

(71) Applicant
Nippon Shinyaku Company Limited (Japan),
14 Kisshoin Nishinosho Monguchicho, Minami-ku,
Kyoto 601, Japan

(72) Inventors
Junichi Yano,
Tadaaki Ogi

(74) Agent and/or Address for Service
Reginald W. Barker & Co., 13 Charterhouse Square,
London EC1M 6BA

(51) INT CL[4]
C12N 15/00

(52) Domestic classification (Edition H):
C3H B4
U1S 1068 C3H

(56) Documents cited

GB A 2125798	EP A2 0035719
GB A 2007670	WO A1 85/01051
EP A1 0090789	WO A1 83/03098

(58) Field of search
C3H
Selected US specifications from IPC sub-class C12N

(54) DNA synthesis

(57) A method of synthesizing long chain DNA, in which blocks having 4 to 8 base sequences are purely chemically ligated by a so-called solid phase method (triester method) using aminated CPG as a carrier.

GB 2 169 605 A

2169605

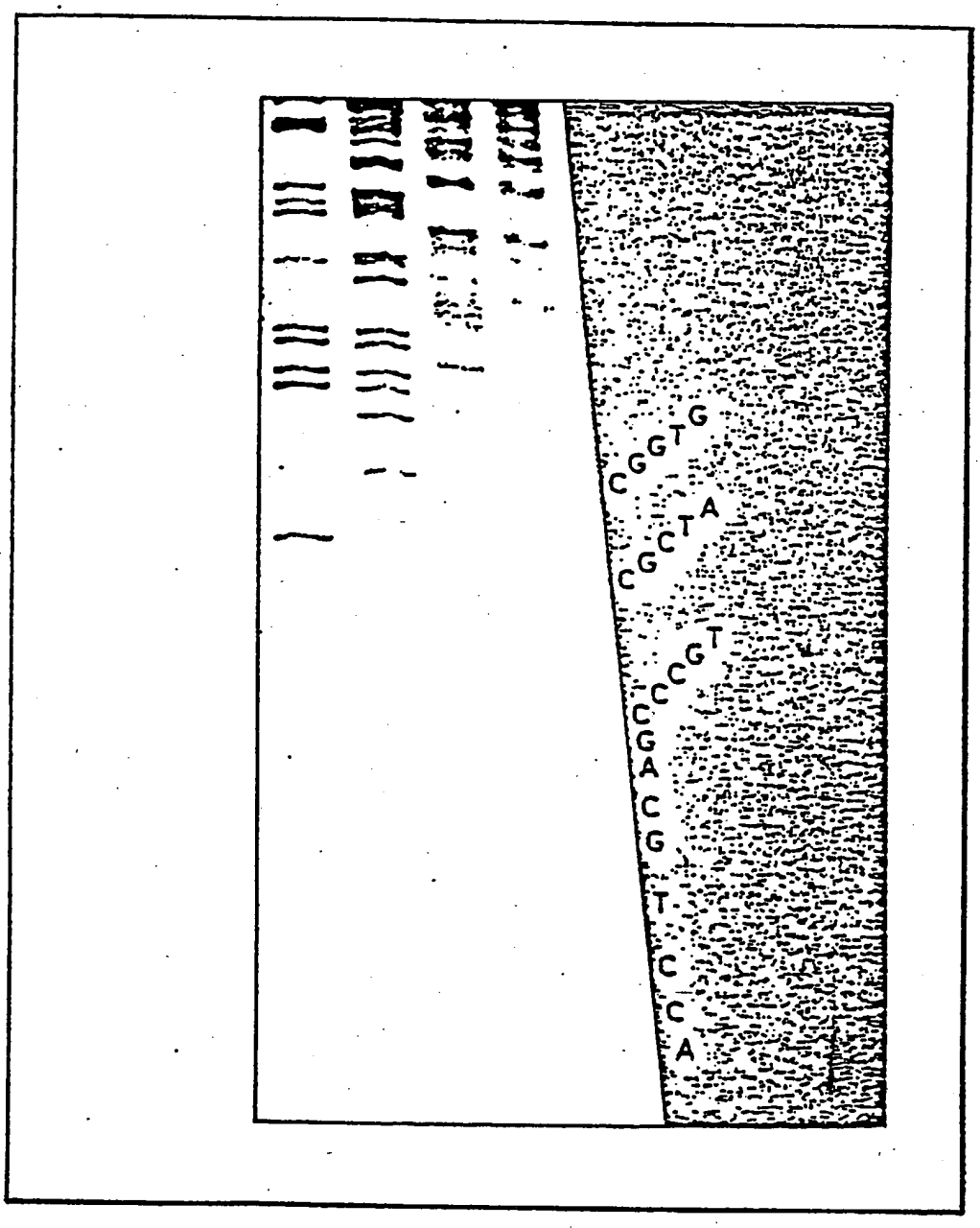


FIG.1

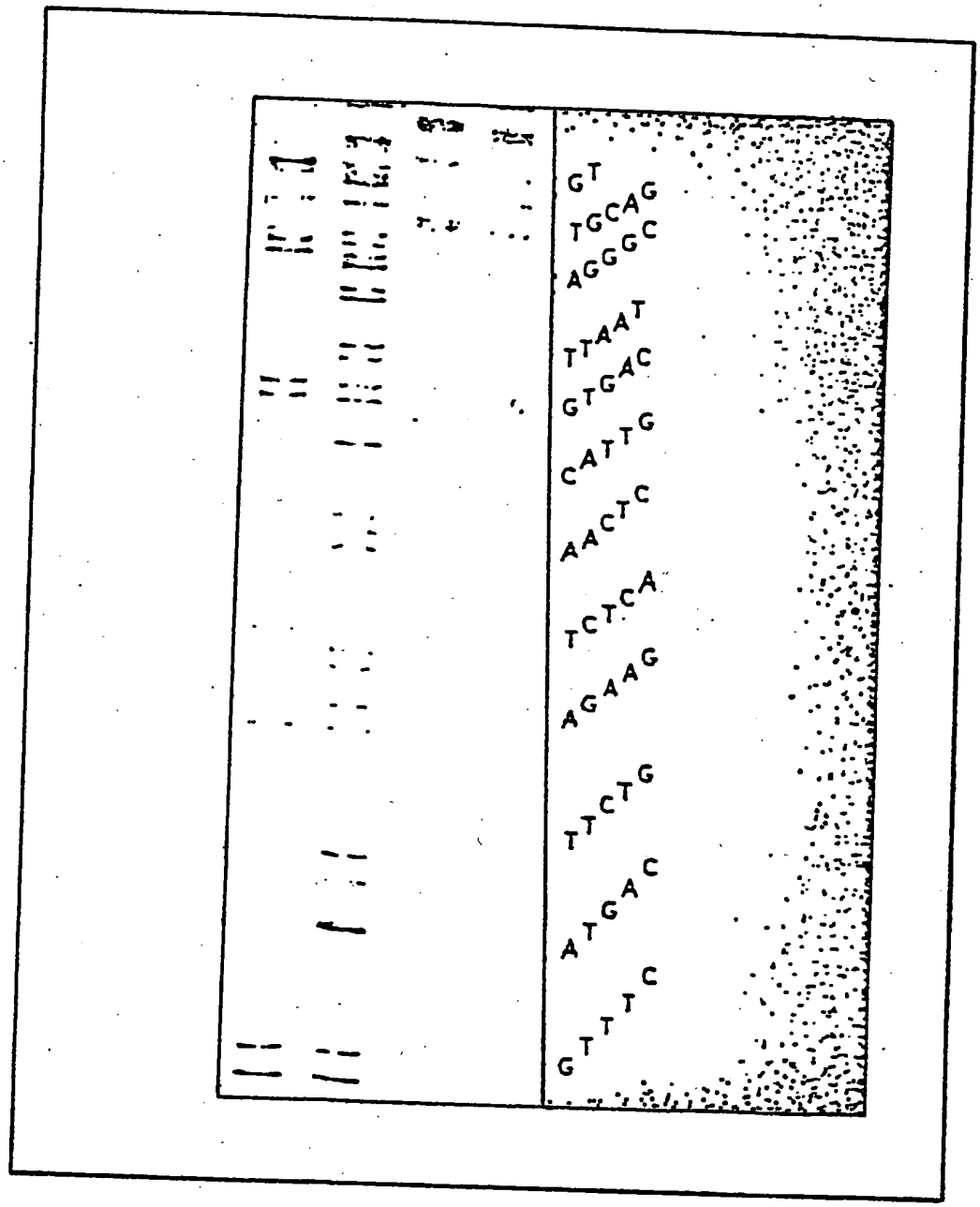

GT
TGCAG
AGGGC
TTAAT
GTGAC
CATTG
AACTC
TCTCA
AGAAG
TTCTG
ATGAC
GTTTC

FIG. 2

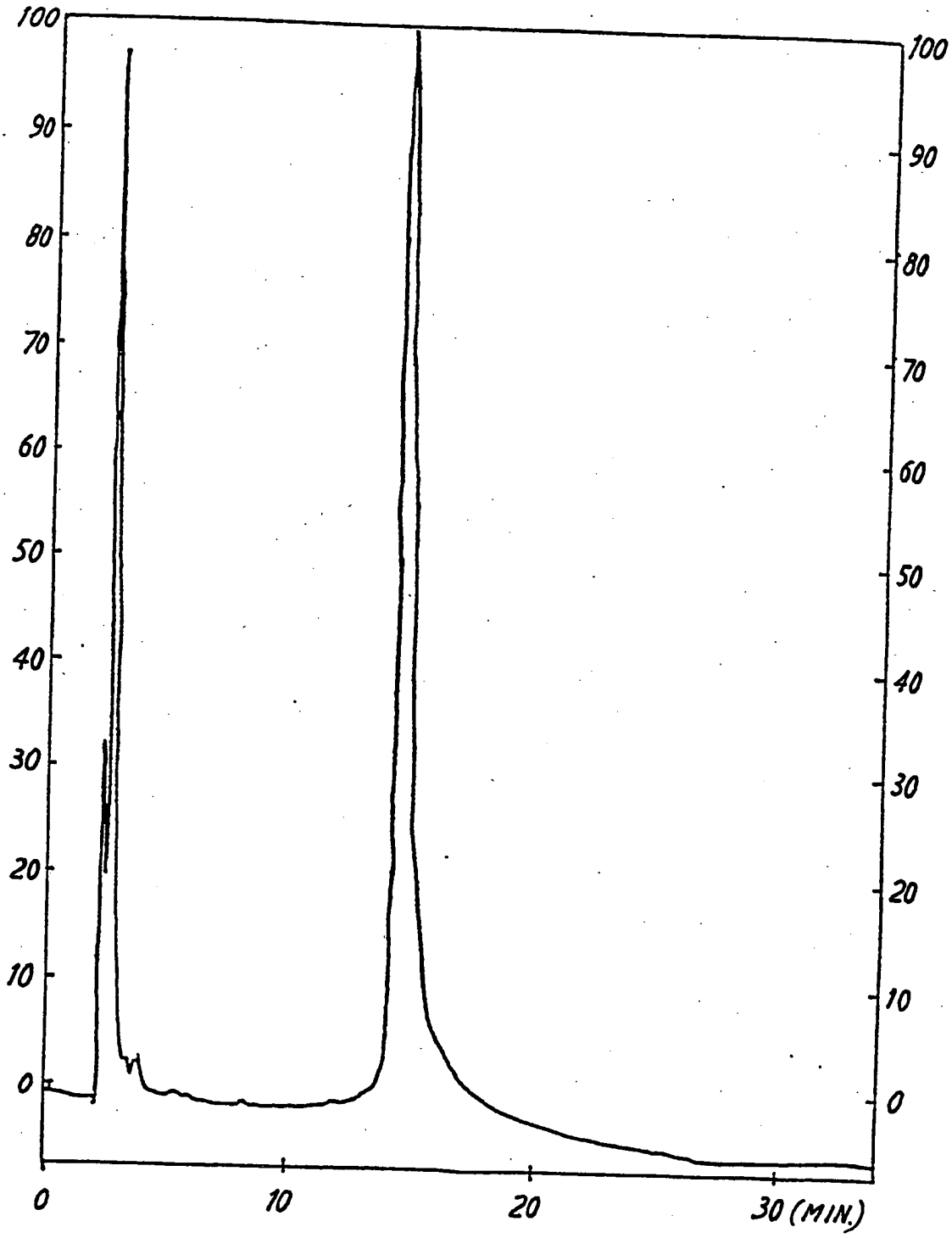


FIG.3

SPECIFICATION

A method of synthesizing long chain DNA

The present invention relates to a novel method of synthesizing long chain DNA carrying information for synthesis of specific proteins and, more particularly, it relates to a method of synthesizing long chain DNA purely chemically, i.e. without the use of enzymes.

It has been known that the synthesis of polypeptides by a gene technological means using synthetic gene is possible by steps of (1) synthesis of structural gene; (2) recombination of the gene into suitable plasmid; (3) transformation of suitable host by the formed chimera plasmid; and (4) obtaining of the desired polypeptide by culturing the transformed substance.

Recently, developments of DNA probe is attracting public attention as a novel means for gene technology. This is a a method of identifying unknown DNA and RNA which is a transcribed product by a hybridization of single stranded DNA and RNA which are known in the art by utilizing properties of DNA and RNA that they form duplex by selecting complimentary substance just like the relation of template and casting. Since very sensitive and prompt identification is possible by utilizing the hybridization method, said method can be applied for diagnosis of precise name of disease by finding specific DNA and RNA in a gene level from blood and cells of patients and pathogenic bacteria. Accordingly, DNA has important value as diagnostic agent by its utilization as DNA probe.

With reference to the above-given DNA as structural gene and that as a source of utilization as DNA probe, it has been known in terms of its nature that the longer the base sequence, the more important as information source and the wider in utilization range to DNA probe. However it has been also known that the longer the base sequence, the more difficult in its synthesis.

Consequently, development of technique in synthesizing long chain DNA by easy manner has been desired.

Conventional method for synthesizing DNA is as follows. Thus, first, comparatively short DNA fragment with 10 to 20 basic residues is chemically synthesized, then they are combined to prepare fragments having total structure of double stranded DNA exhibiting information on desired peptide synthesis, and finally they are combined using an enzyme called DNA ligase.

However, by such a method, only comparatively short fragments with 1 (monomer), 2 (dimer) or 3 (trimer) bases are manufactured prior to block condensation and it is not possible to synthesize long DNA with 80 residues or the like.

In addition, in said method, it is essential to use an enzyme called DNA ligase. Therefore, in synthesizing double stranded DNA as gene, it is necessary that all base sequences constituting double stranded DNA are synthesized at one time. Accordingly, the above method is not so effective in a process of synthesizing double stranded DNA.

The present inventors have continued studies in order to overcome the above technical difficulty and have succeeded in synthesizing DNA with 46 bases or so by utilizing a method called a triester method (among the so-called solid methods) in which 1% polystyrene is used as a support and the compounds of 4 (tetramer) or 5 (pentamer) bases are subjected to a repeated condensation.

Even by such a method, however, the base numbers in the resulting DNA are 50 at the largest and there is still a difficulty in synthesizing DNA with chains of as long as 80 to 150 residues.

(*Problems that the Present Invention Solves*)

In view of the above, the present inventors have further carried out continued studies paying their attention to (1) the synthesis of long chain DNA carrying as much as gene information and (2) the synthesis under more advantageous conditions and finally achieved the present invention.

According to the present invention there is provided:-

A method of synthesizing long chain DNA, characterised in that, blocks having 4 to 8 base sequences are purely chemically ligated by a so-called solid phase method (triester method) using aminated controlled pore glass as a carrier.

The present invention will be further illustrated as hereunder:-

Each block prior to the condensation can be obtained by the conventional way in which each base is subjected to a liquid phase synthesis.

Aminated CPG (controlled pore glass) (cf. Tetrahedron, *24*, 747-750, 1983) used in the present invention is used as a carrier in the solid phase method. To the amino group of this substance is combined deoxythymidine which is changed to 3'-sucinate by usual method. This is used as a carrier for nucle side. Each desired block is extended, on this resin, to the direction of 5'-terminal successively. As to condensation agent, mesitylen sulf nyl-3-nitrotriaz lide (MSNT) can be used, for example. The her by resulting DNA is single stranded and th complimentary strand DNA which is necessary f r preparation of duplet DND can be asily obtained by the similar way. Or such duplet DNA can be very asily obtained by the use of DND polymerase using short fragment (10 b.p. or so) which is complimentary with 3"-terminal region of the resulting single stranded DNA. The fact that DNA polymerase can be used is on of the most advantag ous merits of th present invention that the condensation reaction can b accomplished without the aid of DNA ligase which has been widely used in conventional meth ds.

The resulting duplet DNA is combined to give vector plasmid by the known method, then transformed to bacteria such as *Escherichia coli*, and the strain is cultured to afford desired polypeptid . In the above steps, various gene technological means which have been already established can be applied.

It is possible in accordance with the present invention to synthesize DNA with as long as 80 to 150 residues and, therefore, polypeptides with 15 to 30 amino acids can be synthesized by the known gene technological means. For instance, the following polypeptides can be synthesized. They are growth hormone-release inhibiting factor (Somatostatin, containing 14 amino acids), stomach acid secreting stimulant (Gastrin, containing 17 amino acids), duodenum ulcer remedy (Secretin, containing 27 amino acids), stimulant for secretion of growth hormone, insuline and blood sugar level increase (Glucagon, containing 29 amino acids), morphine like agent (beta-Endorphin, containing 31 amino acids), and hypercalcemia remedy (Calcitonin, containing 32 amino acids), and the like.

In addition, the long chain DNA of the present invention is applied not only for DNA base sequences of structural gene parts but also for the manufacture of general DNA including regulatory sites and specific sequences as well as for long chain DNA probe recognizing their structures. Accordingly the present invention can be positively applied for development of diagnostic agents.

(*Effect of the Invention*)

According to the present invention, long chain DNA can be synthesized simply and in large quantities. The long chain DNA of the present invention can be effectively utilized as (1) gene information source concerning polypeptide synthesis and (2) a source for application of DNA probe in view of gene technology.

Production of DNA has been 0.1 OD (1 OD is equivalent to about 50 micrograms) per one lot at best. However, in accordance with the present invention, it is now possible to manufacture in quantities as large as 30 to 50 OD per lot. Consequently, expansion of utilizable field of long chain DNA as a gene and as a DNA probe can be expected.

(*Examples*)

The present invention is further illustrated by giving examples concerning synthesis of endorphin whose physiological activities such as central nervous analgesic action and endocrine hormone action have been known.

(1) *Synthesis of each block constituting base sequences including endorphin gene.*

Amino acid sequence of endorphins has been known and the DNA base sequence corresponding thereto can be freely selected by referring to a table of coden usage. They are given as hereunder together with their relation between each block constituting DNA base sequences used in the present invention. The upper, middle and lower columns are each block (figures therein are block numbers), base sequence and corresponding amino acid sequence, respectively. Incidentally, restricted enzyme sites are given at both terminals of DNA base sequences. Said sites are used in inserting plasmid.

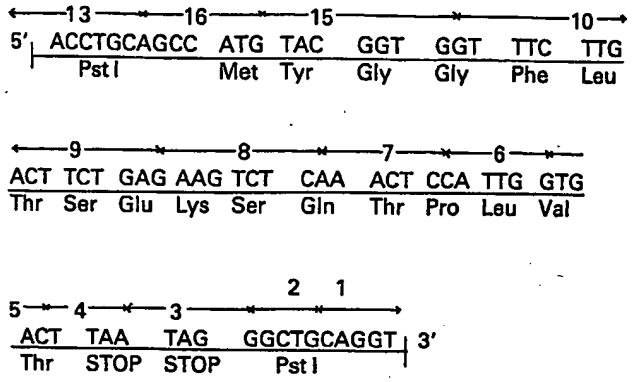
① α – Endorphin

←13→←12→←11→←10→
5' ACCTGCAGCC CGT CGC TAC GGT GGT TTC ATG
Pst I Arg Arg Tyr Gly Gly Phe Met

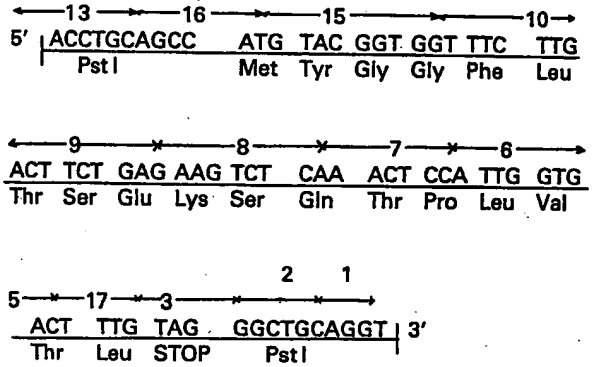
←9→←8→←7→←6→
ACT TCT GAG AAG TCT CAA ACT CCA TTG GTG
Thr Ser Glu Lys Ser Gln Thr Pro Leu Val
2

5→←4→←3→←2→←1→
ACT TAA TAG GGCTGCAGGT 3'
Thr STOP STOP Pst I

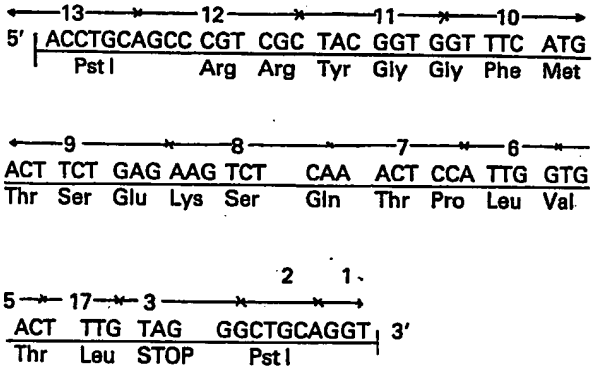
② α-[Leu^5]-Endorphin



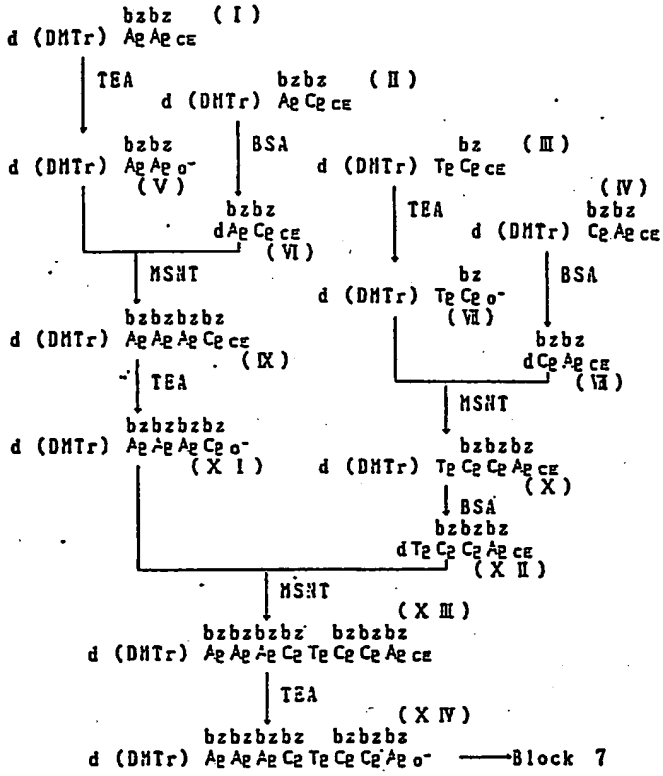
③ γ-[Leu^5]-Endorphin



④ γ-Endorphin



Among the blocks constituting the above endorphin genes, the block 7 was synthesized by the steps as given below.



DMTr: 4,4'-Dimethoxytrityl
$^{bz}_{A}$: N-Benzoyladenosyl
$^{bz}_{C}$: N-Benzoylcytidylyl
T: Tymidylyl
e: o-Chlorophenyl phosphate
BSA: Benzenesulfonic acid
TEA: Triethylamine

Other blocks (1-6, and 8-17) constituting endorphin type genes can be synthesized by similar way. Each yield is given as hereunder.

Blocks	*Base Sequences*	*Yield*
1	CAGG	77
2	GCTG	94
3	TAGG	70
4	TTAA	104
5	TGAC	93
6	TTGG	67
7	AAACTCCA	70
8	GAAGTCTC	62
9	ACTTCTGA	65
10	GTTTCATG	65
11	CTACGGTG	70
12	GCCCGTCA	65
13	ACCTGCA	65
14	GTTTCTTG	70
15	GTACGGTG	67
16	GCCAT	75
17	TTTG	95

(2) *Endorphins genes synthesis:*

alpha-Endorphin gene (deoxy 80 mer) containing restricted enzyme sites was synthesized by a solid phase method as follows.

1. Deoxytymidine CPG resin is washed with CH_2Cl_2/MeOH.
2. Detritylation is conducted with 2% BSA/CH_2Cl_2 (this was conducted repeatedly and promptly until colorization disappears)
3. Subjected to azeotropic drying after substituted with pyridine.
   A solution of each block is added, subjected to azeotropic drying, and MSNT and pyridine for the reaction are added. Allowed to stand at room temperature and washed with pyridine.
4. 0.1M Dimethylaminopyridine/pyridine solution and acetic anhydride are added, allowed to stand at room temperature, and washed with pyridine.

The above operation is conducted repeatedly, for 13 times in total. Average yield of this reaction was 84%. Then the resin is deprotected, at room temperature, with a solution of 0.1M tetramethylguanidine-pyridine aldoxime (cf. C.B. Reese, et al: Tetrahedron Lett., 2727, 1978) in dioxane-water, then washed with pyridine-water, the washing is concentrated *in vacuo*, concentrated ammonia water is added thereto, and the mixture is warmed. Ammonia is evaporated therefrom and a part of the residue is taken using dimethyoxytrityl group as a target to calculate the yield of the final stage.

The residual reaction solution is subjected to a reversed phase (C_{18} silica gel for Prep 500 manufactured by Waters), ion exchange (DEAE-toyopal), and reversed phase (C_{18} silica gel, TSK-Gel 10-20 micrometers) open chromatographies to afford pure alpha-endorphin gene (containing restricted enzyme sites) (– deoxy 80 mer).

Purity was confirmed by HPLC (Nucleosil 300-7 C_{18}) and by electrophoresis and its base sequences were confirmed by Maxam-Gilbert method. The result is given in Figure 1 to Figure 3.

Similarly prepared were alpha-(Leu^5)-endorphin gene (containing restrictive enzyme site) (deoxy 77 mer), gamma-(Leu^5)-endorphin gene (containing restrictive enzyme site) deoxy 77 mer) and gamma-endorphin gene (containing restrictive enzyme site) (deoxy 80 mer).

(3) *Synthesis of duplet DNA and its combination with vector plasmid.*

Each one mole of deoxy 80 mer and synthetic nucleotide primer which is complimentary with 3'-terminal of the former were mixed, heated at 65°C, and cooled to room temperature to anneal the deoxy 80 mer and the primer. Then *E. coli* polymerase I (Klenow fragment) was added by conventional may and made to react at 37°C for 30 minutes so that DNA was made into double stranded.

DNA was recovered as a precipitate in ethanol, made to react at 37°C for 30 minutes using T_4 polynucleotidekinase, and both 5'-terminals of the double stranded DNA were phosphorylated.

Then the vector plasmid pUC 8 DNA was scissored with a restrictive enzyme Pst 1, added to the above double stranded DNA solution, made to react at 16°C overnight with T_4 DNA ligase, and the double stranded d 80 mer DNA was combined with the vector plasmid.

(4) *Cloning of plasmid containing endorphin gene.*

The plasmid prepared as above was transformed into *E. coli* JM 103 strain by conventional way, then selected using a deficiency of beta-galactosidase activity present in the pUC 8 as a target, and plasmid molecules were collected by cloning from the strain.

It has been confirmed that plasmid in which endorphin gene was inserted into the correct orientation and position as desired in accordance with Maxam-Gilbert method.

(5) *Obtaining of endorphins.*

Transformed *E. coli* JM 103 strain was precultured overnight in an LB medium, planted in 2YT medium, and subjected to a shake culture at 37°C.

IPTG was added to the logarithmic productive phase stages (initial, medium and final stages) to make it 0.5mM and synthesis of endorphin was induced. After being induced by IPTG, fused protein was extracted, and analyzed by HPLC whereupon it was found that adequate quantity of protein production was observed (1-5.0 × 10^5 molecules per cell) when induction was applied at the initial stage of logarithmic productive phase.

With reference to natural type alpha-endorphin and gamma-endorphin hacing methionine residue in a molecule, they were treated with trypsin by conventional way. With reference to alpha-(Leu^5)-endorphin and gamma-(Leu^5)-endorphin having leucine residue in place of methionine, they were treated with BrCN. Anyway, each f desired endorphin proteins desired was subjected t a c lumn chromat graphy according t th general purificati n method f proteins whereupon each of them was separat d and purifi d.

The fact that each of the resulting endorphin molecules exhibits desired amino acid sequence was confirmed by the fact that they were identical with the samples already btained by the peptide synthesis by testing with HPLC using a revers phase carri r.

4. *Brief Explanation of Drawings:*

Figure 1 is X-ray autoradi gram showing th result of 20% polyacryamide lectr ph resis f de xy 80 mer containing alpha-endorphin gen synthesized after determinati n by Maxam-Gilbert method.

Figure 2 is X-ray autoradiogram sh wing the result of 8% p lyacrylamide electrophoresis of deoxy 80 mer containing alpha-endorphin gene synthesized after determination by Maxam-Gilbert method.

Figure 3 shows the result of high performance liquid chromatography (Nucleosil 300-7 C_{18}) of deoxy 80 mer containing alpah-endorphin gene synthesized after determination by Maxam-Gilbert method. Ordinate and abscissa show absorbancy and time, respectively. Solvent system used was triethylamine acetate-acetonitrile and the flowing speed was 1.0 ml/min.

CLAIMS

1. A method of synthesizing long chain DNA, characterised in that, blocks having 4 to 8 base sequences are purely chemically ligated by a so-called solid phase method (triester method) using aminated controlled pore glass as a carrier.

2. A method as claimed in Claim 1 in which the blocks having 4 to 8 base sequences are subject to condensation reaction.

3. A method of synthesizing long chain DNA and which is substantially as described herein.

Printed in the UK for HMSO, D8818935, 6/86, 7102.
Published by The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.